Running Title: BabyPy

BabyPy: a brain-age model for infancy, childhood, and adolescence

Francesca Biondo[1,*], Jonathan O'Muircheartaigh[2,3,4], Richard A.I. Bethlehem[5], Jakob Seidlitz[6,7,8,9], Aaron Alexander-Bloch[8,7], Jed T. Elison[10], Viren D'Sa[11], Sean C.L. Deoni[12], Muriel M.K. Bruchhage[13,14,†], James H. Cole[1,15,†]

[1]UCL Hawkes Institute, UCL, London, UK
[2]Department for Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[3]Research Department of Early Life Imaging, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[4]MRC Centre for Neurodevelopmental Disorders, King's College London, London, UK
[5]Department of Psychology, University of Cambridge, Cambridge, UK
[6]Lifespan Brain Institute, The Children's Hospital of Philadelphia and Penn Medicine, Philadelphia, PA, USA
[7]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA
[8]Department of Child and Adolescent Psychiatry and Behavioral Science, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
[9]Institute for Translational Medicine and Therapeutics, University of Pennsylvania, Philadelphia, PA, USA
[10]Institute of Child Development, University of Minnesota, Minneapolis, USA
[11]Rhode Island Hospital, Providence, RI, USA
[12]Bill & Melinda Gates Foundation, Seattle, WA, USA
[13]University of Stavanger, Norway
[14]Stavanger University Hospital, Department of Radiology, Stavanger Medical Imaging Laboratory (SMIL), Stavanger, Norway
[15]Dementia Research Centre, UCL, London, UK

*Corresponding author: f.biondo@ucl.ac.uk
†These authors share senior authorship.

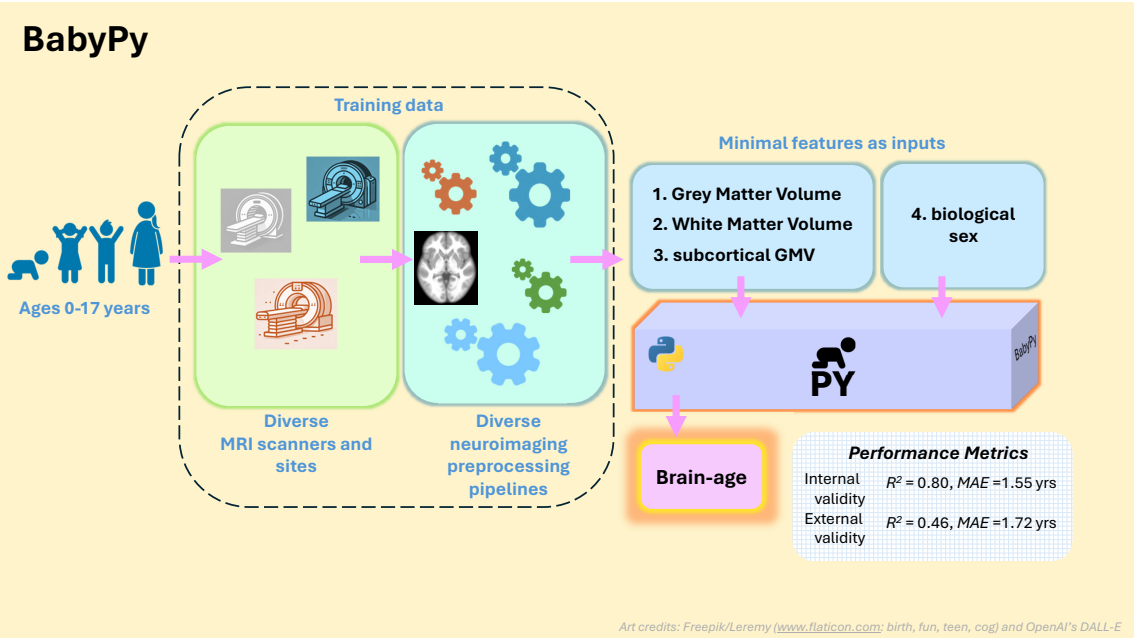
Acknowledgments
We would like to thank the Bill and Melinda Gates Foundation (INV-005774 and INV-047885) for their continuous support of this work and all participants who have made this study possible.

Abstract

Intro: Brain-age models quantify biological ageing by predicting a person's age from neuroimaging data. In early life, brain-age can reflect underlying biological maturity (or immaturity), providing a candidate predictor of typical neurodevelopment versus deviation. Although widely used in adult research, the use of brain-age in early development has been limited due to data availability, heterogeneity and restricted model accessibility. Here, we introduce BabyPy, a shareable brain-age model for individuals aged 0–17 years that achieves accurate predictions despite substantial variability in site, scanner, and preprocessing pipelines. **Methods:** We trained BabyPy on 4,021 structural T1-weighted MRI scans from multi-site datasets (ages 0–17 years). An external test set of 1,143 scans (ages 0–16 years) was used for validation. Coarse neuroimaging features - grey matter volume (GMV), white matter volume (WMV), and subcortical grey matter volume (sGMV) - along with sex, were the model inputs. An ensemble machine learning approach combined Extra Trees Regression, Support Vector Machine, and Multilayer Perceptron base models. Performance was evaluated via 5-fold cross-validation and external testing. **Results:** The ensemble meta-model explained 80% of the variance in age (training set, MAE = 1.55 years) and 46% of the variance in the external test set (MAE = 1.72 years). **Conclusion:** BabyPy is a shareable framework that estimates brain-age across a broad developmental range, removing the need for separate age-specific models. Despite limitations due to data heterogeneity, it demonstrates robust predictive performance and supports cross-study comparisons. Future improvements in data harmonisation will further enhance the utility of generic brain-age models like BabyPy.

Keywords

Brain-age Developmental Neuroimaging Toolbox Machine Learning Neurodevelopment Structural MRI Open Science



Keypoints:

- **Broad developmental coverage:** BabyPy provides a shareable brain-age model for individuals aged 0–17 years, removing the need for separate age-specific models in the primary developmental period.
- **Minimal features, robust performance:** Using just three coarse volumetric measures (GMV, WMV, sGMV) plus sex, BabyPy achieves good predictive accuracy ($R^2 = 0.80$, MAE = 1.55 years in training; $R^2 = 0.46$, MAE = 1.72 years in external validation) across diverse sites, scanners, and preprocessing protocols.
- **Open Science contribution:** BabyPy is freely available as a Python-based toolbox, facilitating easy adoption, reproducibility, and cross-study comparisons in developmental neuroimaging.

1. Introduction

Brain-age is an index of the brain's biological age, providing a simple metric of brain health that is typically condensed into a single value per participant (Cole and Franke, 2017; Gaser et al., 2024), although multiple brain-ages per person are also possible (Smith et al., 2020; Popescu et al., 2021). Brain-age is typically derived by training a model to estimate chronological age based on neuroimaging features from a healthy cohort as a normative reference. This estimated age, or brain-age, has previously been shown to be robust and sensitive to both current and future cognitive and health outcomes in adults (Gaser et al., 2013; Cole and Franke, 2017; Liem et al., 2017; Cole et al., 2018; Kaufmann et al., 2019; Franke and Gaser, 2019; Wang et al., 2019; Cole, 2020; Cole et al., 2020; Smith et al., 2020; Biondo et al., 2022; Leonardsen et al., 2022; Wood et al., 2022; Gaser et al., 2024).

Whilst its applicability to younger cohorts remains relatively under-explored, as barriers to paediatric neuroimaging are overcome and larger datasets become accessible, a growing number of developmental brain-age studies have emerged. Some of these brain-age models utilise across-the-lifespan neuroimaging data (Kaufmann et al., 2019; Bashyam et al., 2020; Bellantuono et al., 2021; Jirsaraie et al., 2023; Leonardsen et al., 2023) and a few focus specifically on neonates, early infancy and fetuses (Hong et al., 2020; Shi et al., 2020; Kawaguchi et al., 2021; Taoudi-Benchekroun et al., 2022; Hu et al., 2023; Fang et al., 2024; Ansari et al., 2024). However, the majority concentrate on parts of the primary developmental period from birth to late adolescence with some extending to early adulthood (Franke et al., 2012; Chung et al., 2018; Zhao et al., 2019; Ball et al., 2021; Brouwer et al., 2021; Cropley et al., 2021; Mendes et al., 2021; Drobinin et al., 2022; Kelly et al., 2022; Lund et al., 2022; Dehestani et al., 2023; Cohen et al., 2023; Mendes et al., 2023; Li et al., 2023; Zhou et al., 2023; Beck et al., 2024; Kozhemiako et al., 2024; Zapaishchykova et al., 2024).

These developmental brain-age models employ a variety of neuroimaging features. Many use parcellations of structural MRI scans derived from segmentation tools such as FreeSurfer (Chung et al., 2018; Zhao et al., 2019; Cropley et al., 2021; Ball et al., 2021; Drobinin et al., 2022; Zhou et al., 2023), grey matter (GM)/white matter (WM) maps via voxel-based morphometry (VBM; Franke et al. (2012); Brouwer et al. (2021); Mendes et al. (2023)) or features derived from saliency-maps(Zapaishchykova et al., 2024). Other studies include diffusion tensor imaging (DTI; Beck et al. (2024); Fang et al. (2024)) or connectivity metrics from structural MRI or resting-state functional MRI (Bellantuono et al., 2021; Lund et al., 2022; Taoudi-Benchekroun et al., 2022; Li et al., 2023), while others use non-MRI neuroimaging data such as EEG (Ansari et al., 2024; Kozhemiako et al., 2024).

There is considerable diversity in brain-age study designs. Many integrate datasets from different sources, including various studies, sites, and scanners, into a single model or multiple models (Franke et al., 2012; Chung et al., 2018; Brouwer et al., 2021; Bellantuono et al., 2021; Ball et al., 2021; Drobinin et al., 2022; Kelly et al., 2022; Zhou et al., 2023; Leonardsen et al., 2023; Dehestani et al., 2023; Cohen et al., 2023; Zapaishchykova et al., 2024), whereas others assign datasets to separate models (Mendes et al., 2021, 2023; Lund et al., 2022). Furthermore, the types of models employed vary widely, with the majority using machine learning algorithms like Gaussian Process Regression (Ball et al., 2021; Kelly et al., 2022); Support/Relevance

Vector Machines or Regression (Franke et al., 2012; Brouwer et al., 2021; Cropley et al., 2021; Zhou et al., 2023; Li et al., 2023; Dehestani et al., 2023); ridge regression (Chung et al., 2018; Cohen et al., 2023); XGBoost (Ball et al., 2021; Drobinin et al., 2022; Beck et al., 2024); SCFN-reg (Leonardsen et al., 2023) and deep learning approaches such as 2D or 3D Convolutional Neural Networks (Bashyam et al., 2020; Kawaguchi et al., 2021; Mendes et al., 2021, 2023; Jirsaraie et al., 2023); Deep Neural Network (Bellantuono et al., 2021; Taoudi-Benchekroun et al., 2022); Gradient Tree Boosting (Kaufmann et al., 2019; Jirsaraie et al., 2023); ResNETs (Hu et al., 2023) and Diffusion models (Zapaishchykova et al., 2024).

Some of these studies apply brain-age as a biomarker by assessing its associations with, or predictions of, cognitive and clinical outcomes. Typically, they do so by utilising brainPAD, the Predicted Age Difference (also referred to by other names such as brain-age gap), calculated as the difference between the estimated brain-age and the participant's chronological age. In early life, brainPAD could reflect underlying biological maturity (or immaturity), providing a potential marker or predictor of neurodevelopmental progress or deviation. The results are mixed: some studies report psychopathology and neurodevelopmental differences, such as ADHD and autism, associated with younger-looking brains (Tunç et al., 2019; Lund et al., 2022; Li et al., 2023); others find the opposite, particularly for psychosis (Chung et al., 2018; Truelove-Hill et al., 2020; Cropley et al., 2021; Drobinin et al., 2022); some report null or small effects (Ball et al., 2021; Mendes et al., 2021); and others show mixed results (Cohen et al., 2023; Zhou et al., 2023; Beck et al., 2024).

Despite evidence supporting the utility of brain-age as a biomarker, developmental brain-age models have limitations regarding generalisability. First, not all models are validated using external test sets, making it unclear whether they generalise to unseen datasets.

Second, the inaccessibility of the models may hinder comparisons across different studies and conditions. For instance, within MRI-based brain-age models, few share their pre-trained models, with some exceptions (Kaufmann et al., 2019; Bashyam et al., 2020; Brouwer et al., 2021; Leonardsen et al., 2022; Drobinin et al., 2022; Jirsaraie et al., 2023; Zapaishchykova et al., 2024). This lack of sharing is likely due to various challenges. For example, while multi-site datasets provide the benefit of larger training sets, they also introduce difficulties in minimising non-biological noise driven by differences in sites, scanners and preprocessing pipelines. Although harmonisation techniques can partially address these issues (Fortin et al., 2018), implementing them in shareable models remains challenging or limited.

Third, to the best of our knowledge, there is currently no developmental brain-age model that encompasses the entire primary developmental period from infancy to adolescence. Lifespan brain-age models (Kaufmann et al., 2019; Bashyam et al., 2020; Jirsaraie et al., 2023; Leonardsen et al., 2023) and others (Zapaishchykova et al., 2024) systematically exclude ages 0 to 3 years, while other models focus exclusively on this infancy period (Hong et al., 2020; Kawaguchi et al., 2021; Hu et al., 2023). This division necessitates the use of separate brain-age models to cover the full range of primary development, potentially introducing comparability issues and complicating longitudinal or cross-sectional developmental studies.

Fourth, a different form of inaccessibility can arise from the specific data types required by the model. For instance, models that rely on less commonly available data, such as raw data rather than derived neuroimaging phenotypes, may have restricted applicability.

In this study, we developed BabyPy, a developmental brain-age model designed to be both shareable and widely adoptable, while maintaining good performance accuracy despite heterogeneity arising from both biological sources, such as developmental variability across ages, and non-biological factors including study and preprocessing differences. BabyPy utilises collaborative multi-site data from individuals aged 0–17 years. We validated the model using an external test set and have made BabyPy, along with its associated tools and scripts, publicly available for use by the research community.

2. Materials and Methods

2.1. Data sources and permissions

Data were obtained from anonymised primary datasets of healthy children, either using existing secondary data research analysis (Bethlehem et al., 2022) via NIH data sharing agreements or through our funding body's agreements (see Supplementary Materials, Table S1, for details). For all children, written informed consent was obtained from their parents or legal guardians in the primary studies. Details on the ethical oversight from local institutions, consent procedures and eligibility criteria can be found in the references provided for each primary dataset in Table S1. Our focus on childhood and adolescence led us to set an age limit of 17, as our external test set did not include data beyond this age.

2.2. Data acquisition and preprocessing

Structural T1-weighted MRI scans were acquired using either 3T or 1.5T MRI scanners. Detailed acquisition parameters for each contributing dataset are provided in Table S1. A flowchart outlining BabyPy model construction is presented in Figure 1.

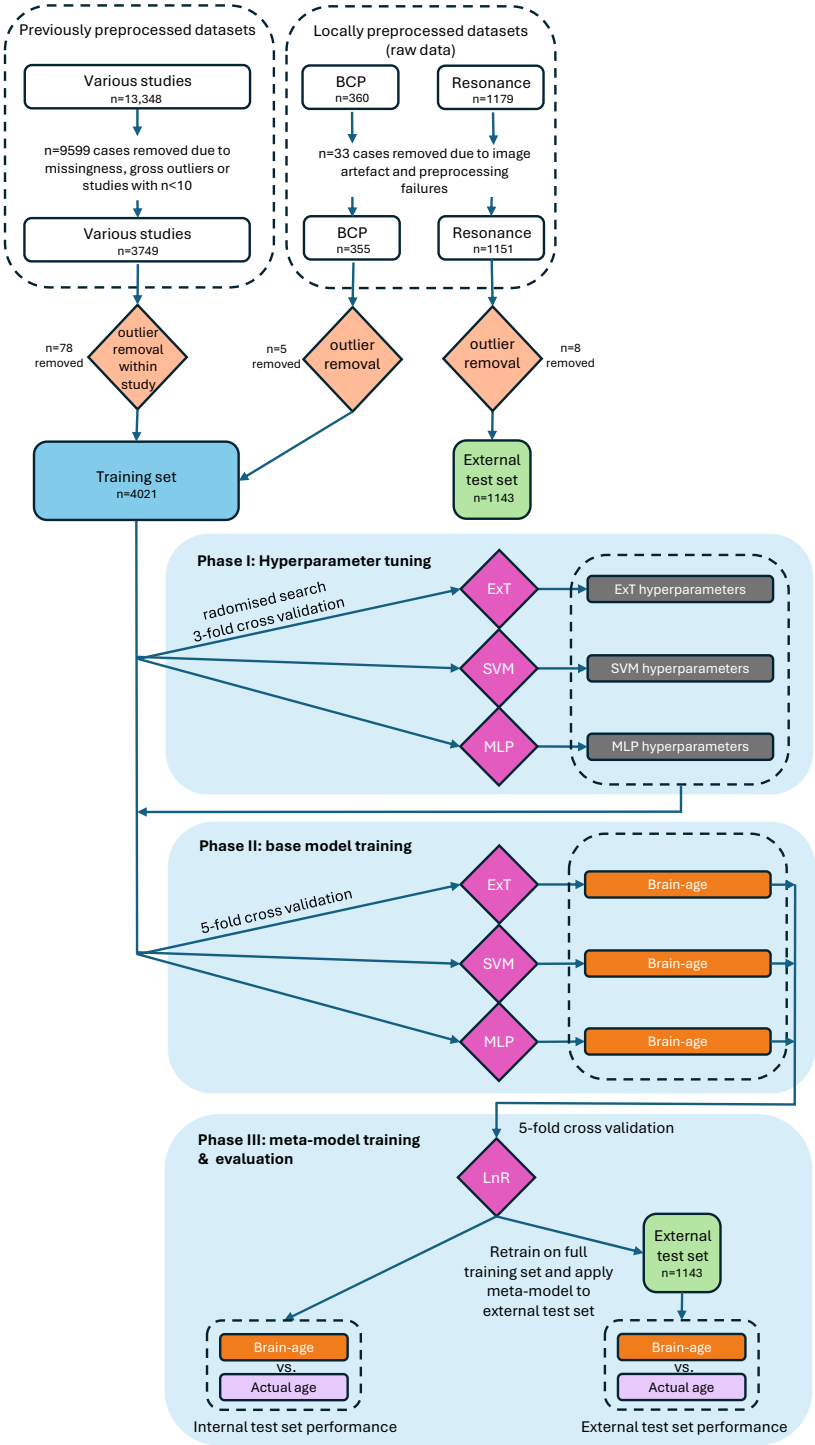


Figure 1: Flowchart of BabyPy model construction. The training set comprised of 13 datasets that had been previously preprocessed (Bethlehem et al., 2022) alongside one locally preprocessed dataset (BCP; Howell et al. (2019)). The external test set consisted of the locally preprocessed Resonance dataset (Bruchhage et al., 2020). Preprocessing steps involved the removal of cases due to gross outliers, studies with a sample size of fewer than 10 participants, missing data, image artefacts, and preprocessing failures. Additional cases were excluded through a custom outlier removal process within each study. In the first phase of model training, hyperparameters for each base model, Extra Trees regression (ExT), Support Vector Machine (SVM), and Multilayer Perceptron (MLP), were optimised using a randomised search with 3-fold cross-validation. The second phase involved training the three base models using 5-fold cross-validation, stratified by sex, incorporating the hyperparameters determined in phase one. In the third phase, an ensemble model was developed by training a meta-model using Linear Regression with the predicted brain-age outputs from phase two, as inputs. Finally, the meta-model was retrained on the entire training set and applied to the external test set. ExT = Extra Trees Regression; SVM = Support Vector Machine; MLP = Multilayer Perceptron, LnR = Linear Regression.

2.2.1. Training set preprocessing

The initial pool of data for training, comprised 13,708 MRI scans from participants aged 0-17 years detailed in Table S1. The preprocessing for almost the entire training set (94%) was conducted previously by collaborators with access to the raw data Bethlehem et al. (2022). Below is a description of their preprocessing methods.

FreeSurfer *recon-all* pipelines were used for segmentation, with the version choice depending on data availability and age. For instance, when T2/FLAIR-weighted scans were available, the combined T1-T2 recon-all pipeline was utilised for enhanced grey-white matter delineation. This resulted in the following 5 types of FreeSurfer pipelines: T1-T2 pipeline (FS6_T1T2), only T1 (FS6_T1), both using FreeSurfer v6.0.142, only T1 using FreeSurfer v5.3 (FS53) and only T1 using Infant FreeSurfer v.1.108 (FSInfant).

The *recon-all* process involves multiple pre-processing steps, including uniformity corrections, intensity normalisation, mapping to Talairach space, skull-stripping, tissue and volume segmentation, and registration. This process generates estimates of GMV, WMV and sGMV, labelled respectively as "Total cortical gray matter volume", "Total cortical white matter volume or Total cerebral white matter volume" and "Subcortical grey matter volume" (inclusive of the thalamus, caudate nucleus, putamen, pallidum, hippocampus, amygdala, and nucleus accumbens area; `https://freesurfer.net/fswiki/SubcorticalSegmentation`).

The remaining 6% of the training set was preprocessed via custom pipelines: Custom1 or Custom2. These were either performed at the primary data site (for Custom1, see details on the HBN dataset via the study reference available in Table S1). For Custom2, this is described in the external test set section below.

For distributions by study and/or preprocessing pipeline see Table 1 and Figure 2 (for extensive details, see Table S1).

2.2.2. External test set preprocessing (Custom2)

Two datasets, Resonance (n=1179; Bruchhage et al. (2020)) and Baby Connectome Project (BCP, n=360; Howell et al. (2019)), were provided in raw, unprocessed format. The Resonance dataset comprises the entire external test set, while the BCP dataset is part of the training set. Both datasets were preprocessed using a custom in-house pipeline described below. Image artefacts or pre-processing failures led to the removal of 28 cases from the Resonance dataset and 5 cases from the BCP dataset.

Image registration followed a two-stage process to align images to a common space executed using the Advanced Normalisation Tools (ANTs) package with symmetric diffeomorphic normalisation (SyN,Avants et al. (2011, 2008); Pieper (2023)).

In the first stage, nonlinear registration was performed to an age-specific T1-weighted template to mitigate the changing T1-weighted contrast in this age-group (Deoni et al., 2012; Sadeghi et al., 2013). There were 11 age-specific templates that mapped to these (natal) age mid-points: 1, 3, 6, 9, 12, 15, 18, 21, 24, 30 and 36 months (for details, see Table S2).

The second stage used a pre-computed spatial transformation from the age-specific template to a common space, specifically the 12-month template, and Montreal Neurological Institute (MNI) space. Registration to MNI space allowed the propagation of an atlas to each child's native space. The atlas used was the Mindboggle Oasis joint fusion atlas (`OASIS-TRT-20_jointfusion_DKT31_CMA_labels_in_MNI` Tustison et al. (2014); Bakker et al. (2015); Klein et al. (2017); RN1 (2024)). Estimates of cortical GMV and sGMV were calculated by summing the relevant regions of interest (see Table S3 and Table S4 for a list).

Label-propagation was used to estimate WMV, due to the variability in resolution and image contrast between Resonance and BCP datasets as well as the pronounced age variability. A 3-class segmentation of the 12-month template, excluding the cerebellum, was performed using FSL FAST. This segmentation achieved estimates of GM, WM, and CSF, which were then propagated back to individual infant and child scans. Total Brain Tissue Volume (TBTV) was calculated by summing GMV and WMV from the 3-class segmentation. WMV was then calculated by subtracting sGMV and GMV derived from the atlas-based segmentations, from TBTV. CSF estimates were not included in our model as extra-axial CSF tends to get eroded to different degrees in brain extraction.

2.3. Data heterogeneity and missingness

To address potential heterogeneity driven by study effects, we initially considered harmonising the training set. However, given our aim to share the model and the current limitations associated with harmonising new, unseen data, we opted not to apply harmonisation in the current analysis. Importantly, the model's performance with harmonised data was comparable to that without harmonisation. Detailed information about the model version with harmonisation can be found in Supplementary Materials, Section 4.4.2.

2.3.1. Missingness

Training set data with the FS53 preprocessing pipeline had missing WMV segmentations and the ABCD dataset using the FS6_T1T2 preprocessing pipeline resulted in gross outliers. Consequently, these 9,242 cases were removed. Additionally, 353 cases were excluded due to missing features, and 4 more cases were removed when we excluded studies with study sizes less than 10, further reducing the previously preprocessed part of the training set to n=3749 scans.

2.3.2. Outlier removal

As part of our quality control, we aimed to remove gross outliers. However, due to the significant volumetric changes associated with age in early development, applying a standard interquartile range (IQR) method across the entire dataset would have led to substantial misclassification. To address this, we employed a custom outlier removal tool that first segments the feature distribution into age bins and then applies the IQR method within each bin. Although an IQR multiplier of 1.5 is typical, we adopted a multiplier of 2 to target only gross outliers. Furthermore, to mitigate bias from study-based differences, the tool was applied within each study rather than across the entire training dataset. After outlier removal, the external test set was reduced from n=1151 to n=1143, and the training set (n=3749

previously preprocessed + n=355 BCP data) was reduced from n=4104 to n=4021. For details on the outlier removal tool, see Section 3 of the Supplementary Materials.

2.3.3. Sample size and repeat scans

The reported sample sizes refer to the number of scans rather than the number of unique individuals. To maximize the available data we retained repeat scans from the same individuals. After applying all data filters, 6.6% of the training set and 72.7% of the external test set scans originated from the same individuals. The higher proportion of repeat scans in the external test set is attributable to its accelerated longitudinal design. Exact numbers are detailed in Table S1.

2.4. Demographics - age and sex

Recognising that early brain development is sensitive to the length of the gestational period, we adjusted natal age (age since birth) by adding the gestational period. This age adjustment is also referred to as postmenstrual age. For participants with unknown gestational age, we assumed a full-term gestational period of 280 days. In this study, all ages are considered adjusted unless explicitly specified as 'natal'.

In this study, sex refers to the biological sex and is categorised as either male or female.

2.5. BabyPy model training

For model training, we adopted an ensemble approach utilising diverse base models. We employed three base models: Extra Trees (ExT), Support Vector Machine (SVM), and a Multilayer Perceptron (MLP).

Python packages were used; *sklearn* for model training and *scipy* for deriving performance metrics. The extracted neuroimaging features: GMV, WMV, sGMV and sex were used as the input features, with age as the target variable. All variables except sex were scaled to unit variance and zero mean before model training. We evaluated performance by comparing the estimated brain-age to the actual age. We used Pearson's R, the explained variance using predictive R^2 (Scheinost et al., 2019), Mean Absolute Error (MAE), and Mean Squared Error (MSE) to assess performance. The model training process consisted of three distinct phases:

2.5.1. Phase 1: hyperparameter tuning

Hyperparameter tuning using scikit-learn's *RandomizedSearchCV* with 3-fold cross-validation and 2000 iterations.

2.5.2. Phase 2: base model training with k-fold cross-validation

Base model training using a 5-fold cross-validation approach, stratified by sex, with the optimised hyperparameters.

2.5.3. Phase 3: meta-model training and evaluation

An ensemble meta-model was developed by training a Linear Regression model (using 5-fold cross-validation, stratified by sex) on the predicted brain-age outputs from the base models. The meta-model was then retrained on the entire training set and applied to both training and external test sets.

2.6. BrainPAD

Brain Predicted Age Difference (BrainPAD) was calculated by subtracting the estimated brain-age from the actual chronological age for each participant. Pearson's R was used as a performance metric.

3. Results

Details for the training and test sets, after completion of all preprocessing and prior to BabyPy model training, are presented in Table 1, with their distributions shown in Figure 2.

Group	N	Mean age (yrs)	SD	Median age	Age range	Male %
training	4021	9.17	4.71	10.08	0.79 - 16.99	55.4%
test	1143	5.22	3.53	4.58	0.94 - 15.77	57.3%

Table 1: Summary of the training and test sets, with mean (SD), median, range in years and sex.

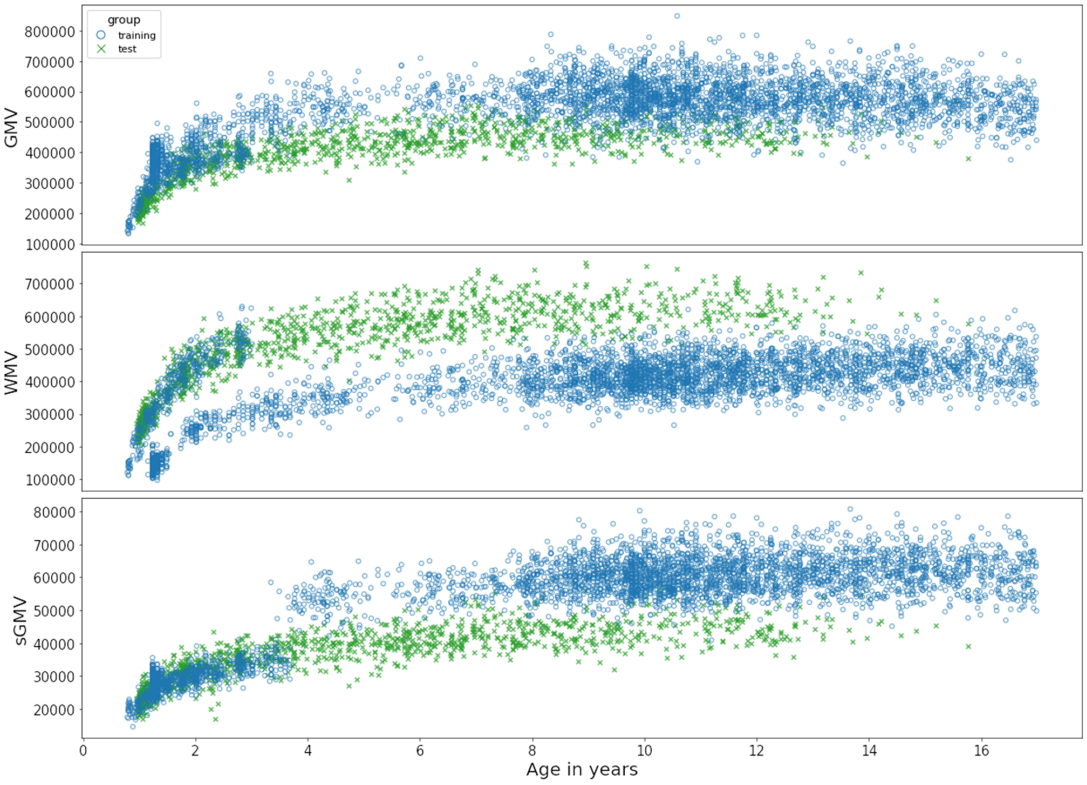


Figure 2: Distribution of neuroimaging features as a function of age across training and external test set categories. The top, middle and bottom scatterplots show the distributions of GMV, WMV and sGMV in mm3, respectively. Blue hollow circles represent the training set, used to train BabyPy. Green crosses represent the external test set.

Using 5-fold cross-validation, in the internal test folds, the meta-model achieved a mean (SD) Pearson's r of 0.89 (0.01), mean R^2 of 0.80 (0.01), an MSE of 4.50 (0.24), and an MAE of 1.55 (0.05). A scatter plot averaged across all 5 folds is presented on the left in Figure 3. Scatter plots for each fold are displayed in Figure S11.

After retraining on the entire training set, the meta-model was evaluated on the external test set, where it attained a Pearson's r of 0.78, R^2 of 0.46, MSE of 6.79, and an MAE of 1.72. A scatter plot for the fully trained meta-model on the external test set is displayed on the right, in Figure 3.

After calculating BrainPAD, we observed negative slopes when plotting these values against chronological age. The Pearson correlation coefficients were $r =$ -0.45 for the training set and $r =$ -0.83 for the test set, indicating a proportional bias linked to age, driven by a regression-to-the-mean effect (see Figure S12).

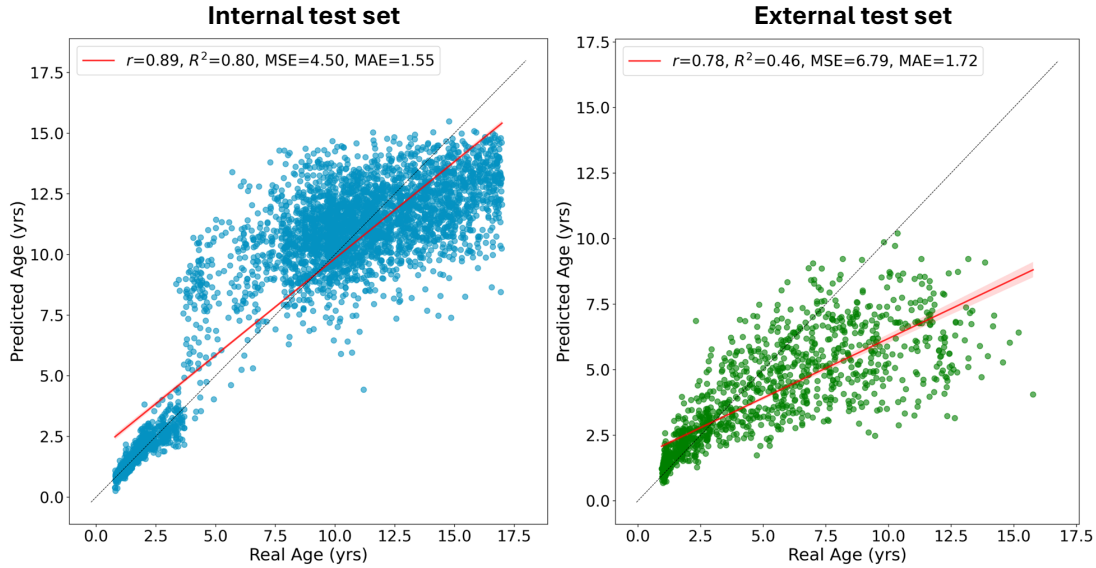


Figure 3: Performance plots for the BabyPy meta-model on the fully trained set. The left scatterplot in blue shows the relationship between real age (years) and predicted brain-age (years) for the internal test set during cross-validation, averaged across all 5 folds. The right scatterplot in green depicts the fully trained meta-model's performance on the external test set. Embedded performance metrics include Pearson's r, predictive R^2, Mean Squared Error (MSE) in years, and Mean Absolute Error (MAE) in years. The diagonal dashed line represents the line of identity, while the red line indicates the line of best fit.

4. Discussion

In this study, we aimed to develop a brain-age model spanning infancy to adolescence that is suitable for public sharing, while maintaining robust predictive performance despite data heterogeneity at many levels, including site, acquisition and preprocessing.

BabyPy achieved a Mean Absolute Error (MAE) of less than 2 years for both cross-validation (1.55) and the external test set (1.72). Performance on this metric is comparable to that of other developmental brain-age models with similar age ranges of 3–21 years, as demonstrated by studies such as Ball et al. (2021) with cross-validation MAE = 1.75, Chung et al. (2018) with cross-validation MAE = 1.69 and external validation MAE = 1.41 in an older 12–21 year cohort, Kelly et al. (2022) with cross-validation MAE = 1.72 and external validation MAE = 1.92 and MAE = 1.53 at age 7 and age 13, respectively and, Zapaishchykova et al. (2024) trained using a 3-30 years age range and with two external validation tests achieving MAEs of 2.78 and 1.97.

However, directly comparing brain-age models across published studies is challenging due to varying study objectives, the influence of multiple factors related to data and model types (Jirsaraie et al., 2023) as well as the difference in the type of metrics reported. Importantly, the distribution of the test set ages often differs between studies, meaning the MAEs are not comparable. More common reporting of predictive R^2, a recommended metric for predictive models like brain-age (Scheinost et al., 2019), would increase comparability.

Age is a key source of biological heterogeneity, given the rapid and significant morphological and functional changes occurring during development (Giedd et al.,

1999; Silk and Wood, 2011). In addition, paediatric image acquisition inherently involves significant noise and signal variability, particularly in GM and WM tissue segmentation during the first years of life (Deoni et al., 2012). Consequently, the age range used for both training and testing influences model performance accuracy. For example, in the studies cited directly above, the ones reporting external validation were on held-out sets with narrower and older age ranges than their 3-21-year-old training set. It is hard to ascertain what the performance would be if their test set encompassed the entire training age range instead.

While other models either exclude or focus exclusively on the infancy period, our model spans the entire primary developmental period (0-17 years). This broad scope may result in higher MAE compared to models which exclude infancy and use narrower age bands. For example, (Cohen et al., 2023) used a training set of 5–17 years, achieving a cross-validation MAE of 1.41 years. Mendes et al. (2023) utilised a training set aged 7–20 years and reported an external validation MAE of 1.56 years. Zhao et al. (2019) included participants aged 5–18 years with an external validation MAE of 1.42 years. Cropley et al. (2021) trained on a cohort aged 8–21 years and reported a cross-validation MAE of 1.49 years. Two studies using the large ABCD dataset, spanning the narrow range of 9–13 years, reported cross-validation MAEs of 0.63 years (Dehestani et al., 2023) and 0.47 years (Mendes et al., 2023) albeit with an external validation MAE of 2.17 years (Mendes et al., 2023).

However, excluding infancy is not always accompanied by a lower MAE. For instance, (Li et al., 2023) employed a training set aged 5–21 years with a leave-one-out cross-validation MAE of 2.28 years and (Drobinin et al., 2022) utilised a training set aged 9–19 years with an external validation MAE of 1.86 years. Two lifespan brain-age models (Kaufmann et al., 2019; Bashyam et al., 2020; Jirsaraie et al., 2023), spanning approximately 3-96 years of age, demonstrated relatively poorer performance when tested on a sample within the primary developmental phase (e.g., 11-17 years), with mean absolute errors (MAEs) ranging from 2.8-23 years (Jirsaraie et al., 2023). Retraining these models by restricting the sample size to ages 8-22 significantly improved performance with their 'DBN' model's MAE decreasing from 4-21.7 to 1.7-4.7.

Equally hard to compare are studies focusing on infancy, as the MAE is expected to scale to the training set range. For example, the following studies applied minimally processed data with deep learning approaches in infancy, achieving MAEs as low as 2.2 months using a ResNet-18 for 0–3 years (Hu et al., 2023), 67.6 days with 3D CNN for 0–5 years (Hong et al., 2020)) and 1.9 months using CNN for 0–24 months (Kawaguchi et al., 2021).

We observe that BabyPy's R^2 decreases significantly from training to the external test set (from 0.8 to 0.46). This reduction likely reflects the heterogeneity of the data. In other words, there is a trade-off: a model that is generalisable across diverse ages, sites, scanners, and preprocessing pipelines tends to sacrifice accuracy when tailored to any specific combination of these factors. Nonetheless, the relationship between data heterogeneity (driven by both biological and non-biological sources), and model accuracy and generalisability is not straightforward. For instance, some studies argue that homogeneity within the training set may improve out-of-sample performance by encouraging models to focus on robust features (Mendes et al., 2021,

2023). Others have found that diversity in preprocessing pipelines — such as using multiple versions of FreeSurfer rather than a single version — can improve brain-age performance (Korbmacher et al., 2024).

When we attempted to address some of the non-biological heterogeneity through study-based harmonisation, as presented in the supplementary materials, we observed small to modest improvements. In the external test set, R^2 increased from 0.46 to 0.58 and the MAE decreased from 1.72 to 1.64. Overall, although harmonisation led to some improvements, these gains remain modest, as similarly reported by others (Kelly et al., 2022). However, these metrics alone may not capture the full cost-benefit trade-offs.

BabyPy exhibits a regression-to-the-mean effect, where predicted brain-age estimates are biased towards the mean age of the training set. Specifically, younger participants tend to have older-looking brain-age estimates, while older participants show younger-looking estimates. This bias was clearly reflected in the training set and more extensively in the external test set with BrainPAD versus age correlation coefficients of -0.45 and -0.83, respectively. This phenomenon is typical of brain-age models and can be mitigated through various post-hoc correction methods (Smith et al., 2019; de Lange and Cole, 2020). We have not implemented these corrections in BabyPy, and hence we recommend users consider them in downstream analyses involving brainPAD, for example, by incorporating both linear and non-linear forms of age as covariates (see Section 4.3 of the Supplementary).

Overall, despite the variability in age, sites, scanners, and preprocessing methods, BabyPy demonstrates performance accuracy within the typical-to-high range of existing developmental brain-age models, which is promising. However, its utility value will become more apparent as it is applied to various targets and contexts. To this end, we highlight BabyPy's adoptability. Firstly, it is user-friendly and requires only two main steps: volume extraction and model application. Because it has been trained using data from a wide array of sites, protocols, and preprocessing steps, BabyPy is more likely robust and compatible with real-world conditions of unseen data characterised by heterogeneity (Ball et al., 2021; Bethlehem et al., 2022). Additionally, BabyPy is low-dimensional, employing coarse parcellations - GMV, WMV and sGMV - which enhance compatibility compared to models using large-feature inputs including voxel-based. Coarse parcellations facilitate noise smoothing across different preprocessing pipelines and hence reduce dependency on any specific segmentation pipelines, making the model more flexible across different datasets. This is particularly relevant in the context of research platforms, where tabular rather than raw data is more readily shared. In contrast, in clinical settings, models that minimise preprocessing may be preferable (Jirsaraie et al., 2023). Overall, we hope this high level of adoptability will encourage comparisons across various targets and studies, thereby providing a useful tool to better research and reconcile the mixed brain-age findings reported in younger cohorts (Tunç et al., 2019; Cropley et al., 2021; Mendes et al., 2021; Drobinin et al., 2022; Lund et al., 2022; Li et al., 2023; Zhou et al., 2023).

4.1. Limitations and future directions

This study presents several limitations and suggests directions for future research.

Firstly, as previously discussed, there is a trade-off between accuracy and robustness across diverse datasets, which constrains BabyPy's accuracy for any specific combination of characteristics of any given dataset. A primary focus for future work should be the improvement of harmonisation techniques. This is not only in terms of reducing non-biological noise, but also retaining compatibility with unseen data — similar to those achieved within different, but related methodologies (see Generalised Additive Models for Location, Scale, and Shape (GAMLS) Bethlehem et al. (2022), transfer learning, and Hierarchical Bayesian Regression in normative modelling Gaiser et al. (2024)).

Another limitation concerns the trade-off between generic and specific models. While a generic model, like BabyPy, can account for a broad range of differences, it may inadvertently obscure important sub-group distinctions that could benefit from specialised models. For example, sex differences (Brouwer et al., 2021) and developmental timing, such as key stages in late childhood and adolescence, where incorporating pubertal status could be particularly informative (Brouwer et al., 2021; Dehestani et al., 2023) or critical transition points during infancy and early childhood (Lund et al., 2022; Bethlehem et al., 2022). However, determining the precise boundaries for such distinctions remains challenging. Moreover, although BabyPy does not currently account for differences in head size, this limitation may reduce its generalisability to low- and middle-income countries, where growth stunting is more common (Black et al., 2017) and can influence brain development.

Lastly, while brain-age serves as a convenient metric for summarising overall brain health, this simplicity may overlook the complex interactions of increasing and decreasing brain tissue types during development (Giedd et al., 1999; Silk and Wood, 2011; Bethlehem et al., 2022). Alternative normative modelling approaches that employ multiple models, each targeting specific brain features, might better capture these nuanced differences, offering a more detailed understanding (Smith et al., 2020; Bethlehem et al., 2022; Gaiser et al., 2024; Ho et al., 2024). Ultimately, the suitability of BabyPy, other brain-age models, and normative modelling approaches will depend on the specific research objectives.

5. Conclusion

BabyPy is a novel brain-age model offering a robust and shareable framework for individuals aged from birth to adolescence. Despite limitations related to data heterogeneity, BabyPy demonstrates robust predictive performance. BabyPy can facilitate cross-study comparisons, potentially shedding light on brain health patterns across the primary developmental period without the need for separate specific models. Future improvements, particularly in prospective data harmonisation, will further enhance the utility of generic brain-age models like BabyPy.

CRediT Author Statement

Francesca Biondo: Conceptualisation, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Software, Validation, Visualisation, Writing – Original Draft.
Jonathan O'Muircheartaigh: Data Curation, Formal Analysis (neuroimaging preprocessing), Project Administration, Writing – Review & Editing.
Richard A.I. Bethlehem: Data Curation, Formal Analysis (neuroimaging preprocessing), Writing – Review & Editing.
Jakob Seidlitz: Data Curation, Formal Analysis (neuroimaging preprocessing), Writing – Review & Editing.
Aaron Alexander-Bloch: Data Curation, Formal Analysis (neuroimaging preprocessing), Writing – Review & Editing.
Jed T. Elison: Data Curation, Funding Acquisition, Writing – Review & Editing.
Viren D'Sa: Data Curation, Funding Acquisition.
Sean C.L. Deoni: Data Curation, Funding Acquisition, Writing – Review & Editing.
Muriel M.K. Bruchhage: Data Curation, Funding Acquisition,Project Administration, Writing – Review & Editing.
James H. Cole: Conceptualisation, Supervision, Funding Acquisition, Project Administration, Writing – Review & Editing.

Conflict of interest disclosure

All authors declare no competing interests except: JS, RAIB and AA-B hold shares in and JS and RAIB are directors of Centile Bioscience; JHC is shareholder/advisor for BrainKey and Claritas HealthTech.

Ethics approval

This research used multiple datasets, each overseen by its own local institutional ethics committee. Details about these approvals are provided in the references for each dataset in Table S1. Access to these data was granted via NIH DAR ID 10481 and DAR ID 10683, and through a data sharing agreement between UCL (PI: Prof. James Cole) and the Bill & Melinda Gates Foundation (2021).

Funding

This research study was funded by the Bill and Melinda Gates Foundation (INV-005774 and INV-047885).

Data and Toolbox Availability

The original data are not publicly available; however, access can be requested through each source study's application procedures managed by each source study, as detailed in Table S1.
We provide three resources on GitHub: (1) the pre-trained BabyPy model, accompanied by a guide for applying it to new data, (2) the code used to train the BabyPy model, and (3) the outlier removal tool. [NOTE: these will be made available soon]

References

, 2024. URL: https://mindboggle.info/data.html, doi:https://mindboggle.info/data.html.

Ansari, A., Pillay, K., Arasteh, E., Dereymaeker, A., Mellado, G.S., Jansen, K., Winkler, A.M., Naulaers, G., Bhatt, A., Huffel, S.V., Hartley, C., Vos, M.D., Slater, R., Baxter, L., 2024. Resting state electroencephalographic brain activity in neonates can predict age and is indicative of neurodevelopmental outcome. Clinical Neurophysiology 163, 226–235. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85194056884&doi=10.1016%2fj.clinph.2024.05.002&partnerID=40&md5=b82099ef47eba8527d39e50b589ecb1bhttps://pmc.ncbi.nlm.nih.gov/articles/PMC11250083/pdf/main.pdf, doi:10.1016/j.clinph.2024.05.002.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12, 26–41. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2276735/pdf/nihms41509.pdf, doi:10.1016/j.media.2007.06.004.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ants similarity metric performance in brain image registration. NeuroImage 54, 2033–2044. URL: https://pubmed.ncbi.nlm.nih.gov/20851191https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3065962/, doi:10.1016/j.neuroimage.2010.09.025.

Bakker, R., Tiesinga, P., Kötter, R., 2015. The scalable brain atlas: Instant web-based access to public brain atlases and related content. Neuroinformatics 13, 353–366. URL: https://doi.org/10.1007/s12021-014-9258-x, doi:10.1007/s12021-014-9258-x.

Ball, G., Kelly, C.E., Beare, R., Seal, M.L., 2021. Individual variation underlying brain age estimates in typical development. NeuroImage 235, 118036. URL: https://www.sciencedirect.com/science/article/pii/S105381192100313X, doi:https://doi.org/10.1016/j.neuroimage.2021.118036.

Bashyam, V.M., Erus, G., Doshi, J., Habes, M., Nasrallah, I.M., Truelove-Hill, M., Srinivasan, D., Mamourian, L., Pomponio, R., Fan, Y., Launer, L.J., Masters, C.L., Maruff, P., Zhuo, C., Völzke, H., Johnson, S.C., Fripp, J., Koutsouleris, N., Satterthwaite, T.D., Wolf, D., Gur, R.E., Gur, R.C., Morris, J., Albert, M.S., Grabe, H.J., Resnick, S., Bryan, R.N., Wolk, D.A., Shou, H., Davatzikos, C., on behalf of the ISTAGING Consortium, the Preclinical Alzheimer's disease Consortium, A., studies, C., 2020. Mri signatures of brain age and disease over the lifespan based on a deep brain network and 14468 individuals worldwide. Brain 143, 2312–2324. URL: https://doi.org/10.1093/brain/awaa160, doi:10.1093/brain/awaa160.

Beck, D., Whitmore, L., MacSweeney, N., Brieant, A., Karl, V., de Lange, A.M.G., Westlye, L.T., Mills, K.L., Tamnes, C.K., 2024. Dimensions of early-life adversity are differentially associated with patterns of delayed and accelerated brain maturation. Biological Psychiatry URL: https://www.scopus.com/

inward/record.uri?eid=2-s2.0-85205387714&doi=10.1016%2fj.biopsych.2024.07.019&partnerID=40&md5=f18de0d1a3c70c9fe7fc717efe55cce2, doi:10.1016/j.biopsych.2024.07.019.

Bellantuono, L., Marzano, L., La Rocca, M., Duncan, D., Lombardi, A., Maggipinto, T., Monaco, A., Tangaro, S., Amoroso, N., Bellotti, R., 2021. Predicting brain age with complex networks: From adolescence to adulthood. NeuroImage 225. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85094176283&doi=10.1016%2fj.neuroimage.2020.117458&partnerID=40&md5=a9809291342611a01d9828627f4b89a8, doi:10.1016/j.neuroimage.2020.117458.

Bethlehem, R.A.I., Seidlitz, J., White, S.R., Vogel, J.W., Anderson, K.M., Adamson, C., Adler, S., Alexopoulos, G.S., Anagnostou, E., Areces-Gonzalez, A., Astle, D.E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S.A., Benegal, V., Beyer, F., Blangero, J., Blesa Cábez, M., Boardman, J.P., Borzage, M., Bosch-Bayard, J.F., Bourke, N., Calhoun, V.D., Chakravarty, M.M., Chen, C., Chertavian, C., Chetelat, G., Chong, Y.S., Cole, J.H., Corvin, A., Costantino, M., Courchesne, E., Crivello, F., Cropley, V.L., Crosbie, J., Crossley, N., Delarue, M., Delorme, R., Desrivieres, S., Devenyi, G.A., Di Biase, M.A., Dolan, R., Donald, K.A., Donohoe, G., Dunlop, K., Edwards, A.D., Elison, J.T., Ellis, C.T., Elman, J.A., Eyler, L., Fair, D.A., Feczko, E., Fletcher, P.C., Fonagy, P., Franz, C.E., Galan-Garcia, L., Gholipour, A., Giedd, J., Gilmore, J.H., Glahn, D.C., Goodyer, I.M., Grant, P.E., Groenewold, N.A., Gunning, F.M., Gur, R.E., Gur, R.C., Hammill, C.F., Hansson, O., Hedden, T., Heinz, A., Henson, R.N., Heuer, K., Hoare, J., Holla, B., Holmes, A.J., Holt, R., Huang, H., Im, K., Ipser, J., Jack, C.R., Jackowski, A.P., Jia, T., Johnson, K.A., Jones, P.B., Jones, D.T., Kahn, R.S., Karlsson, H., Karlsson, L., Kawashima, R., Kelley, E.A., Kern, S., Kim, K.W., Kitzbichler, M.G., Kremen, W.S., Lalonde, F., Landeau, B., et al., 2022. Brain charts for the human lifespan. Nature 604, 525–533. URL: https://doi.org/10.1038/s41586-022-04554-y, doi:10.1038/s41586-022-04554-y.

Biondo, F., Jewell, A., Pritchard, M., Aarsland, D., Steves, C.J., Mueller, C., Cole, J.H., 2022. Brain-age is associated with progression to dementia in memory clinic patients. NeuroImage: Clinical 36. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137295286&doi=10.1016%2fj.nicl.2022.103175&partnerID=40&md5=5b010f6ff1bfc19596d2a782c3675d24, doi:10.1016/j.nicl.2022.103175.

Black, M.M., Walker, S.P., Fernald, L.C.H., Andersen, C.T., DiGirolamo, A.M., Lu, C., McCoy, D.C., Fink, G., Shawar, Y.R., Shiffman, J., Devercelli, A.E., Wodon, Q.T., Vargas-Barón, E., Grantham-McGregor, S., 2017. Early childhood development coming of age: science through the life course. The Lancet 389, 77–90. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85002291911&doi=10.1016%2fS0140-6736%2816%2931389-7&partnerID=40&md5=290c9b4211457101822b848f8ff412a7, doi:10.1016/S0140-6736(16)31389-7.

Brouwer, R.M., Schutte, J., Janssen, R., Boomsma, D.I., Hulshoff Pol, H.E., Schnack, H.G., 2021. The speed of development of adolescent brain age

depends on sex and is genetically determined. Cerebral Cortex 31, 1296–1306. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85100280504&doi=10.1093%2fcercor%2fbhaa296&partnerID=40&md5=f812ce5ff546c90a43f159a748e1d868, doi:10.1093/cercor/bhaa296.

Bruchhage, M.M.K., Ngo, G.C., Schneider, N., D'Sa, V., Deoni, S.C.L., 2020. Functional connectivity correlates of infant and early childhood cognitive development. Brain Structure and Function 225, 669–681. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85079536701&doi=10.1007%2fs00429-020-02027-4&partnerID=40&md5=5abfa99911455f706c9206a9f9ee5558, doi:10.1007/s00429-020-02027-4.

Chung, Y., Addington, J., Bearden, C.E., Cadenhead, K., Cornblatt, B., Mathalon, D.H., McGlashan, T., Perkins, D., Seidman, L.J., Tsuang, M., Walker, E., Woods, S.W., McEwen, S., Van Erp, T.G.M., Cannon, T.D., 2018. Use of machine learning to determine deviance in neuroanatomical maturity associated with future psychosis in youths at clinically high risk. JAMA Psychiatry 75, 960–968. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053032715&doi=10.1001%2fjamapsychiatry.2018.1543&partnerID=40&md5=f6fb639ff2e3aad60ec082fc6ce553ab, doi:10.1001/jamapsychiatry.2018.1543.

Cohen, J.W., Ramphal, B., DeSerisy, M., Zhao, Y., Pagliaccio, D., Colcombe, S., Milham, M.P., Margolis, A.E., 2023. Relative brain age is associated with socioeconomic status and anxiety/depression problems in youth. Developmental Psychology 60, 199–209. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85182588211&doi=10.1037%2fdev0001593&partnerID=40&md5=657c99b827f96845aa09189f9ede6f09https://pmc.ncbi.nlm.nih.gov/articles/PMC10993304/pdf/nihms-1969814.pdf, doi:10.1037/dev0001593.

Cole, J.H., 2020. Multimodality neuroimaging brain-age in uk biobank: relationship to biomedical, lifestyle, and cognitive factors. Neurobiology of Aging 92, 34–42. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084034649&doi=10.1016%2fj.neurobiolaging.2020.03.014&partnerID=40&md5=9bcbb3591dde7dbd72ce665168f8da5b, doi:10.1016/j.neurobiolaging.2020.03.014.

Cole, J.H., Franke, K., 2017. Predicting age using neuroimaging: Innovative brain ageing biomarkers. Trends in Neurosciences 40, 681–690. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85034704875&doi=10.1016%2fj.tins.2017.10.001&partnerID=40&md5=b0e9410869f16e8772a37d030a98ca95, doi:10.1016/j.tins.2017.10.001.

Cole, J.H., Raffel, J., Friede, T., Eshaghi, A., Brownlee, W.J., Chard, D., De Stefano, N., Enzinger, C., Pirpamer, L., Filippi, M., Gasperini, C., Rocca, M.A., Rovira, A., Ruggieri, S., Sastre-Garriga, J., Stromillo, M.L., Uitdehaag, B.M.J., Vrenken, H., Barkhof, F., Nicholas, R., Ciccarelli, O., group, M.s., 2020. Longitudinal assessment of multiple sclerosis with the brain-age paradigm. Annals of Neurology 88, 93–105. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.25746, doi:https://doi.org/10.1002/ana.25746.

Cole, J.H., Ritchie, S.J., Bastin, M.E., Valdés Hernández, M.C., Muñoz Maniega, S., Royle, N., Corley, J., Pattie, A., Harris, S.E., Zhang, Q., Wray, N.R.,

Redmond, P., Marioni, R.E., Starr, J.M., Cox, S.R., Wardlaw, J.M., Sharp, D.J., Deary, I.J., 2018. Brain age predicts mortality. Molecular Psychiatry 23, 1385–1392. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85029716076&doi=10.1038%2fmp.2017.62&partnerID=40&md5=eab8c9fcea7dbd6f422364c2e008a313, doi:10.1038/mp.2017.62.

Cropley, V.L., Tian, Y., Fernando, K., Mansour L, S., Pantelis, C., Cocchi, L., Zalesky, A., 2021. Brain-predicted age associates with psychopathology dimensions in youths. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 6, 410–419. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85091713177&doi=10.1016%2fj.bpsc.2020.07.014&partnerID=40&md5=29aa42e8c4e91827bb76ffcaba8a990e, doi:10.1016/j.bpsc.2020.07.014.

Dehestani, N., Whittle, S., Vijayakumar, N., Silk, T.J., 2023. Developmental brain changes during puberty and associations with mental health problems. Developmental Cognitive Neuroscience 60. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85150051019&doi=10.1016%2fj.dcn.2023.101227&partnerID=40&md5=6002e13d6626e302546087c25a847bf9https://www.sciencedirect.com/science/article/pii/S1878929323000324?via%3Dihub, doi:10.1016/j.dcn.2023.101227.

Deoni, S.C.L., Dean, Douglas C., r., O'Muircheartaigh, J., Dirks, H., Jerskey, B.A., 2012. Investigating white matter development in infancy and early childhood using myelin water faction and relaxation time mapping. NeuroImage 63, 1038–1053. URL: https://pubmed.ncbi.nlm.nih.gov/22884937https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3711836/, doi:10.1016/j.neuroimage.2012.07.037.

Drobinin, V., Van Gestel, H., Helmick, C.A., Schmidt, M.H., Bowen, C.V., Uher, R., 2022. The developmental brain age is associated with adversity, depression, and functional outcomes among adolescents. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 7, 406–414. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85123841648&doi=10.1016%2fj.bpsc.2021.09.004&partnerID=40&md5=9b5f4aa6b616d3fb2cfb4691c72f58d0, doi:10.1016/j.bpsc.2021.09.004.

Fang, Z., Pan, N., Liu, S., Li, H., Pan, M., Zhang, J., Li, Z., Liu, M., Ge, X., 2024. Comparative analysis of brain age prediction using structural and diffusion mris in neonates. NeuroImage 299. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85202851419&doi=10.1016%2fj.neuroimage.2024.120815&partnerID=40&md5=b3c6f65a2f5e835ab717294cbfd3c4fdhttps://www.sciencedirect.com/science/article/pii/S1053811924003124?via%3Dihub, doi:10.1016/j.neuroimage.2024.120815.

Fortin, J.P., Cullen, N., Sheline, Y.I., Taylor, W.D., Aselcioglu, I., Cook, P.A., Adams, P., Cooper, C., Fava, M., McGrath, P.J., McInnis, M., Phillips, M.L., Trivedi, M.H., Weissman, M.M., Shinohara, R.T., 2018. Harmonization of cortical thickness measurements across scanners and sites. NeuroImage 167, 104–120. URL: https://www.sciencedirect.com/science/article/pii/S105381191730931X, doi:https://doi.org/10.1016/j.neuroimage.2017.11.024.

Franke, K., Gaser, C., 2019. Ten years of brainage as a neuroimaging biomarker of brain aging: What insights have we gained? Frontiers in Neurology 10. doi:10.3389/fneur.2019.00789.

Franke, K., Luders, E., May, A., Wilke, M., Gaser, C., 2012. Brain maturation: Predicting individual brainage in children and adolescents using structural mri. NeuroImage 63, 1305–1312. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84866152087&doi=10.1016%2fj.neuroimage.2012.08.001&partnerID=40&md5=54be59484f7354a71871434167c465a1, doi:10.1016/j.neuroimage.2012.08.001.

Gaiser, C., Berthet, P., Kia, S.M., Frens, M.A., Beckmann, C.F., Muetzel, R.L., Marquand, A.F., 2024. Estimating cortical thickness trajectories in children across different scanners using transfer learning from normative models. Human Brain Mapping 45. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85184221868&doi=10.1002%2fhbm.26565&partnerID=40&md5=123f097721cd2193ed8e416f4414ba3b, doi:10.1002/hbm.26565.

Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N., Sauer, H., 2013. Brainage in mild cognitive impaired patients: Predicting the conversion to alzheimer's disease. PLoS ONE 8. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84879516484&doi=10.1371%2fjournal.pone.0067346&partnerID=40&md5=93d054db9a4ebc253928cd30164abb9a, doi:10.1371/journal.pone.0067346.

Gaser, C., Kalc, P., Cole, J.H., 2024. A perspective on brain-age estimation and its clinical promise. Nature Computational Science URL: https://doi.org/10.1038/s43588-024-00659-8, doi:10.1038/s43588-024-00659-8.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: A longitudinal mri study [2]. Nature Neuroscience 2, 861–863. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0033305929&doi=10.1038%2f13158&partnerID=40&md5=5b89fbebca5f3d10e7847dcb9eef3f7d, doi:https://doi.org/10.1038/13158.

Ho, N.C.W., Bethlehem, R.A.I., Seidlitz, J., Nogovitsyn, N., Metzak, P., Ballester, P.L., Hassel, S., Rotzinger, S., Poppenk, J., Lam, R.W., Taylor, V.H., Milev, R., Adamson, C., Adler, S., Alexander-Bloch, A.F., Anagnostou, E., Anderson, K.M., Areces-Gonzalez, A., Astle, D.E., Auyeung, B., Ayub, M., Bae, J.B., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S.A., Benegal, V., Bethlehem, R.A.I., Beyer, F., Blangero, J., Cábez, M.B., Boardman, J.P., Borzage, M., Bosch-Bayard, J.F., Bourke, N., Bullmore, E.T., Calhoun, V.D., Chakravarty, M.M., Chen, C., Chertavian, C., Chetelat, G., Chong, Y.S., Corvin, A., Costantino, M., Courchesne, E., Crivello, F., Cropley, V.L., Crosbie, J., Crossley, N., Delarue, M., Delorme, R., Desrivieres, S., Devenyi, G., Di Biase, M.A., Dolan, R., Donald, K.A., Donohoe, G., Dorfschmidt, L., Dunlop, K., Edwards, A.D., Elison, J.T., Ellis, C.T., Elman, J.A., Eyler, L., Fair, D.A., Fletcher, P.C., Fonagy, P., Franz, C.E., Galan-Garcia, L., Gholipour, A., Giedd, J., Gilmore, J.H., Glahn, D.C., Goodyer, I.M., Grant, P.E., Groenewold, N.A., Gudapati, S., Gunning, F.M., Gur, R.E., Gur, R.C., Hammill, C.F., Hansson, O., Hedden, T., Heinz, A., Henson, R.N., Heuer, K., Hoare, J., Holla, B., Holmes, A.J., Huang, H.,

Ipser, J., Jack, C.R., Jackowski, A.P., Jia, T., Jones, D.T., Jones, P.B., Kahn, R.S., Karlsson, H., Karlsson, L., Kawashima, R., et al., 2024. Atypical brain aging and its association with working memory performance in major depressive disorder. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 9, 786–799. URL: https://www.sciencedirect.com/science/article/pii/S2451902224001101, doi:https://doi.org/10.1016/j.bpsc.2024.04.008.

Hong, J., Feng, Z., Wang, S.H., Peet, A., Zhang, Y.D., Sun, Y., Yang, M., 2020. Brain age prediction of children using routine brain mr images via deep learning. Frontiers in Neurology 11. URL: https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2020.584682, doi:10.3389/fneur.2020.584682.

Howell, B.R., Styner, M.A., Gao, W., Yap, P.T., Wang, L., Baluyot, K., Yacoub, E., Chen, G., Potts, T., Salzwedel, A., Li, G., Gilmore, J.H., Piven, J., Smith, J.K., Shen, D., Ugurbil, K., Zhu, H., Lin, W., Elison, J.T., 2019. The unc/umn baby connectome project (bcp): An overview of the study design and protocol development. NeuroImage 185, 891–905. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85044638537&doi=10.1016%2fj.neuroimage.2018.03.049&partnerID=40&md5=e68ea837fcccae4342fa6155e8d5b673, doi:10.1016/j.neuroimage.2018.03.049.

Hu, L., Wan, Q., Huang, L., Tang, J., Huang, S., Chen, X., Bai, X., Kong, L., Deng, J., Liang, H., Liu, G., Liu, H., Lu, L., 2023. Mri-based brain age prediction model for children under 3 years old using deep residual network. Brain Structure and Function 228, 1771–1784. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85168620983&doi=10.1007%2fs00429-023-02686-z&partnerID=40&md5=3a72f4fa44f65baaea8e9d1a63aeffb7https://link.springer.com/article/10.1007/s00429-023-02686-z, doi:10.1007/s00429-023-02686-z.

Jirsaraie, R.J., Kaufmann, T., Bashyam, V., Erus, G., Luby, J.L., Westlye, L.T., Davatzikos, C., Barch, D.M., Sotiras, A., 2023. Benchmarking the generalizability of brain age models: Challenges posed by scanner variance and prediction bias. Human Brain Mapping 44, 1118–1128. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141616648&doi=10.1002%2fhbm.26144&partnerID=40&md5=fbb35ae8ee286676d9272b74beb06d3bhttps://pmc.ncbi.nlm.nih.gov/articles/PMC9875922/pdf/HBM-44-1118.pdf, doi:10.1002/hbm.26144.

Kaufmann, T., van der Meer, D., Doan, N.T., Schwarz, E., Lund, M.J., Agartz, I., Alnæs, D., Barch, D.M., Baur-Streubel, R., Bertolino, A., Bettella, F., Beyer, M.K., Bøen, E., Borgwardt, S., Brandt, C.L., Buitelaar, J., Celius, E.G., Cervenka, S., Conzelmann, A., Córdova-Palomera, A., Dale, A.M., de Quervain, D.J.F., Carlo, P.D., Djurovic, S., Dørum, E.S., Eisenacher, S., Elvsåshagen, T., Espeseth, T., Fatouros-Bergman, H., Flyckt, L., Franke, B., Frei, O., Haatveit, B., Håberg, A.K., Harbo, H.F., Hartman, C.A., Heslenfeld, D., Hoekstra, P.J., Høgestøl, E.A., Jernigan, T.L., Jonassen, R., Jönsson, E.G., Farde, L., Flyckt, L., Engberg, G., Erhardt, S., Fatouros-Bergman, H., Cervenka, S., Schwieler, L., Piehl, F., Agartz, I., Collste, K., Victorsson, P., Malmqvist, A., Hedberg, M., Orhan, F., Kirsch, P., Kłoszewska, I., Kolskår, K.K., Landrø, N.I., Hellard, S.L., Lesch, K.P., Lovestone, S., Lundervold, A., Lundervold, A.J., Maglanoc, L.A., Malt, U.F., Mecocci, P., Melle, I.,

Meyer-Lindenberg, A., Moberget, T., Norbom, L.B., Nordvik, J.E., Nyberg, L., Oosterlaan, J., Papalino, M., Papassotiropoulos, A., Pauli, P., Pergola, G., Persson, K., Richard, G., Rokicki, J., Sanders, A.M., Selbæk, G., Shadrin, A.A., Smeland, O.B., Soininen, H., Sowa, P., Steen, V.M., Tsolaki, M., Ulrichsen, K.M., Vellas, B., Wang, L., Westman, E., Ziegler, G.C., Zink, M., Andreassen, O.A., Westlye, L.T., 2019. Common brain disorders are associated with heritable patterns of apparent aging of the brain. Nature Neuroscience 22, 1617–1623. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072625058&doi=10.1038%2fs41593-019-0471-7&partnerID=40&md5=8db86f1ec7a3a259eda27d8319587eb3, doi:10.1038/s41593-019-0471-7.

Kawaguchi, M., Kidokoro, H., Ito, R., Shiraki, A., Suzuki, T., Maki, Y., Tanaka, M., Sakaguchi, Y., Yamamoto, H., Takahashi, Y., Naganawa, S., Natsume, J., 2021. Age estimates from brain magnetic resonance images of children younger than two years of age using deep learning. Magnetic Resonance Imaging 79, 38–44. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85102894255&doi=10.1016%2fj.mri.2021.03.004&partnerID=40&md5=aa050c8f76a8a5e83aa66b1374fc14b4, doi:10.1016/j.mri.2021.03.004.

Kelly, C., Ball, G., Matthews, L.G., Cheong, J.L.Y., Doyle, L.W., Inder, T.E., Thompson, D.K., Anderson, P.J., 2022. Investigating brain structural maturation in children and adolescents born very preterm using the brain age framework. NeuroImage 247, 118828. URL: https://www.sciencedirect.com/science/article/pii/S1053811921010995, doi:https://doi.org/10.1016/j.neuroimage.2021.118828.

Klein, A., Ghosh, S.S., Bao, F.S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., Keshavan, A., 2017. Mindboggling morphometry of human brains. PLoS Comput Biol 13, e1005350. doi:10.1371/journal.pcbi.1005350.

Korbmacher, M., Westlye, L.T., Maximov, I.I., 2024. Freesurfer version-shuffling can enhance brain age predictions. NeuroImage: Reports 4, 100214. URL: https://www.sciencedirect.com/science/article/pii/S2666956024000205, doi:https://doi.org/10.1016/j.ynirp.2024.100214.

Kozhemiako, N., Buckley, A.W., Chervin, R.D., Redline, S., Purcell, S.M., 2024. Mapping neurodevelopment with sleep macro- and micro-architecture across multiple pediatric populations. NeuroImage: Clinical 41. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85180996987&doi=10.1016%2fj.nicl.2023.103552&partnerID=40&md5=b57d521c50703334ad59a20640b26b5ehttps://www.sciencedirect.com/science/article/pii/S2213158223002437?via%3Dihub, doi:10.1016/j.nicl.2023.103552.

de Lange, A.M.G., Cole, J.H., 2020. Commentary: Correction procedures in brain-age prediction. NeuroImage: Clinical 26. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85080074130&doi=10.1016%2fj.nicl.2020.102229&partnerID=40&md5=dcc23f03ad7458b7836a647124fbc826, doi:10.1016/j.nicl.2020.102229.

Leonardsen, E.H., Peng, H., Kaufmann, T., Agartz, I., Andreassen, O.A., Celius, E.G., Espeseth, T., Harbo, H.F., Høgestøl, E.A., Lange, A.M.d., Marquand,

A.F., Vidal-Piñeiro, D., Roe, J.M., Selbæk, G., Sørensen, , Smith, S.M., Westlye, L.T., Wolfers, T., Wang, Y., 2022. Deep neural networks learn general and clinically relevant representations of the ageing brain. NeuroImage 256, 119210. URL: https://www.sciencedirect.com/science/article/pii/S1053811922003342, doi:https://doi.org/10.1016/j.neuroimage.2022.119210.

Leonardsen, E.H., Vidal-Piñeiro, D., Roe, J.M., Frei, O., Shadrin, A.A., Iakunchykova, O., de Lange, A.M.G., Kaufmann, T., Taschler, B., Smith, S.M., Andreassen, O.A., Wolfers, T., Westlye, L.T., Wang, Y., 2023. Genetic architecture of brain age and its causal relations with brain and mental disorders. Molecular Psychiatry 28, 3111–3120. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85159109386&doi=10.1038%2fs41380-023-02087-y&partnerID=40&md5=cb5522657d46868f8f09d3ef054c8098https://www.nature.com/articles/s41380-023-02087-y.pdf, doi:10.1038/s41380-023-02087-y.

Li, Y., Zhou, Z., Zhang, Y., Ai, H., Liu, M., Liu, J., Wang, L., Qiu, J., Rachel Han, Z., Zhang, Z., Luo, Y.J., Xu, P., 2023. Brain development mediates the relationship between self-reported poor parental monitoring and adolescent anxiety. NeuroImage: Clinical 40. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85173548111&doi=10.1016%2fj.nicl.2023.103514&partnerID=40&md5=82c4a55c47c037b00c22adb3c243ab8dhttps://www.sciencedirect.com/science/article/pii/S221315822300205X?via%3Dihub, doi:10.1016/j.nicl.2023.103514.

Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J.M., Lampe, L., Rahim, M., Abraham, A., Craddock, R.C., Riedel-Heller, S., Luck, T., Loeffler, M., Schroeter, M.L., Witte, A.V., Villringer, A., Margulies, D.S., 2017. Predicting brain-age from multimodal imaging data captures cognitive impairment. NeuroImage 148, 179–188. URL: https://www.sciencedirect.com/science/article/pii/S1053811916306103, doi:https://doi.org/10.1016/j.neuroimage.2016.11.005.

Lund, M.J., Alnæs, D., de Lange, A.M.G., Andreassen, O.A., Westlye, L.T., Kaufmann, T., 2022. Brain age prediction using fmri network coupling in youths and associations with psychiatric symptoms. NeuroImage: Clinical 33. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85121627678&doi=10.1016%2fj.nicl.2021.102921&partnerID=40&md5=0f408089aafbca1eb2f4c2c5b4bf2846, doi:10.1016/j.nicl.2021.102921.

Mendes, S.L., Pinaya, W.H.L., Pan, P., Sato, J.R., 2021. Estimating gender and age from brain structural mri of children and adolescents: A 3d convolutional neural network multitask learning model. Computational Intelligence and Neuroscience 2021. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85107684771&doi=10.1155%2f2021%2f5550914&partnerID=40&md5=eaebb36362238243c681136571b7eeee, doi:10.1155/2021/5550914.

Mendes, S.L., Pinaya, W.H.L., Pan, P.M., Jackowski, A.P., Bressan, R.A., Sato, J.R., 2023. Generalizability of 3d cnn models for age estimation in diverse youth populations using structural mri. Scientific Reports 13. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.

0-85157959681&doi=10.1038%2fs41598-023-33920-7&partnerID=40&md5=0bf5b02bd8e74d11983a8e9282774a69https://www.nature.com/articles/s41598-023-33920-7.pdf, doi:10.1038/s41598-023-33920-7.

Pieper, S., 2023. Advanced normalization tools (ants) URL: http://stnava.github.io/ANTs/.

Popescu, S.G., Glocker, B., Sharp, D.J., Cole, J.H., 2021. A u-net model of local brain-age. bioRxiv , 2021.01.26.428243URL: http://biorxiv.org/content/early/2021/01/27/2021.01.26.428243.abstract, doi:10.1101/2021.01.26.428243.

Sadeghi, N., Prastawa, M., Fletcher, P.T., Wolff, J., Gilmore, J.H., Gerig, G., 2013. Regional characterization of longitudinal dt-mri to study white matter maturation of the early developing brain. NeuroImage 68, 236–247. URL: https://pubmed.ncbi.nlm.nih.gov/23235270https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3693970/, doi:10.1016/j.neuroimage.2012.11.040.

Scheinost, D., Noble, S., Horien, C., Greene, A.S., Lake, E.M., Salehi, M., Gao, S., Shen, X., O'Connor, D., Barron, D.S., Yip, S.W., Rosenberg, M.D., Constable, R.T., 2019. Ten simple rules for predictive modeling of individual differences in neuroimaging. NeuroImage 193, 35–45. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062830717&doi=10.1016%2fj.neuroimage.2019.02.057&partnerID=40&md5=a67a24abe680e513f7a410cb587ac53b, doi:10.1016/j.neuroimage.2019.02.057.

Shi, W., Yan, G., Li, Y., Li, H., Liu, T., Sun, C., Wang, G., Zhang, Y., Zou, Y., Wu, D., 2020. Fetal brain age estimation and anomaly detection using attention-based deep ensembles with uncertainty. NeuroImage 223. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85090220167&doi=10.1016%2fj.neuroimage.2020.117316&partnerID=40&md5=c471111caed101562e60aba039ebbd0f, doi:10.1016/j.neuroimage.2020.117316.

Silk, T.J., Wood, A.G., 2011. Lessons about neurodevelopment from anatomical magnetic resonance imaging. Journal of Developmental Behavioral Pediatrics 32. URL: https://journals.lww.com/jrnldbp/fulltext/2011/02000/lessons_about_neurodevelopment_from_anatomical.13.aspx.

Smith, S.M., Elliott, L.T., Alfaro-Almagro, F., McCarthy, P., Nichols, T.E., Douaud, G., Miller, K.L., 2020. Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. eLife 9, e52677. URL: https://doi.org/10.7554/eLife.52677, doi:10.7554/eLife.52677.

Smith, S.M., Vidaurre, D., Alfaro-Almagro, F., Nichols, T.E., Miller, K.L., 2019. Estimation of brain age delta from brain imaging. NeuroImage 200, 528–539. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068781606&doi=10.1016%2fj.neuroimage.2019.06.017&partnerID=40&md5=f01c255e501b208d31a150be43d9e360, doi:10.1016/j.neuroimage.2019.06.017.

Taoudi-Benchekroun, Y., Christiaens, D., Grigorescu, I., Gale-Grant, O., Schuh, A., Pietsch, M., Chew, A., Harper, N., Falconer, S., Poppe, T., Hughes, E., Hutter, J., Price, A.N., Tournier, J.D., Cordero-Grande, L., Counsell, S.J., Rueckert, D., Arichi, T., Hajnal, J.V., Edwards, A.D., Deprez, M., Batalle, D., 2022. Predicting age and clinical risk from the neonatal connectome. NeuroImage 257, 119319. URL: https://www.sciencedirect.com/science/article/pii/S1053811922004384, doi:https://doi.org/10.1016/j.neuroimage.2022.119319.

Truelove-Hill, M., Erus, G., Bashyam, V., Varol, E., Sako, C., Gur, R.C., Gur, R.E., Koutsouleris, N., Zhuo, C., Fan, Y., Wolf, D.H., Satterthwaite, T.D., Davatzikos, C., 2020. A multidimensional neural maturation index reveals reproducible developmental patterns in children and adolescents. Journal of Neuroscience 40, 1265–1275. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85079077113&doi=10.1523%2fJNEUROSCI.2092-19.2019&partnerID=40&md5=01d6caa52c3d4920df0d3dd3da0fb389, doi:10.1523/JNEUROSCI.2092-19.2019.

Tunç, B., Yankowitz, L.D., Parker, D., Alappatt, J.A., Pandey, J., Schultz, R.T., Verma, R., 2019. Deviation from normative brain development is associated with symptom severity in autism spectrum disorder. Molecular Autism 10. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85076345563&doi=10.1186%2fs13229-019-0301-5&partnerID=40&md5=7afc406c5f1b270a0974038137bfbab6, doi:10.1186/s13229-019-0301-5.

Tustison, N.J., Cook, P.A., Klein, A., Song, G., Das, S.R., Duda, J.T., Kandel, B.M., van Strien, N., Stone, J.R., Gee, J.C., Avants, B.B., 2014. Large-scale evaluation of ants and freesurfer cortical thickness measurements. NeuroImage 99, 166–179. URL: https://www.sciencedirect.com/science/article/pii/S1053811914004091, doi:https://doi.org/10.1016/j.neuroimage.2014.05.044.

Wang, J., Knol, M.J., Tiulpin, A., Dubost, F., de Bruijne, M., Vernooij, M.W., Adams, H.H.H., Ikram, M.A., Niessen, W.J., Roshchupkin, G.V., 2019. Gray matter age prediction as a biomarker for risk of dementia. Proceedings of the National Academy of Sciences 116, 21213. URL: http://www.pnas.org/content/116/42/21213.abstract, doi:10.1073/pnas.1902376116.

Wood, D.A., Kafiabadi, S., Busaidi, A.A., Guilhem, E., Montvila, A., Lynch, J., Townend, M., Agarwal, S., Mazumder, A., Barker, G.J., Ourselin, S., Cole, J.H., Booth, T.C., 2022. Accurate brain-age models for routine clinical mri examinations. NeuroImage 249, 118871. URL: https://www.sciencedirect.com/science/article/pii/S1053811922000015, doi:https://doi.org/10.1016/j.neuroimage.2022.118871.

Zapaishchykova, A., Tak, D., Ye, Z., Liu, K.X., Likitlersuang, J., Vajapeyam, S., Chopra, R.B., Seidlitz, J., Bethlehem, R.A., Mak, R.H., Mueller, S., Haas-Kogan, D.A., Poussaint, T.Y., Aerts, H.J., Kann, B.H., 2024. Diffusion deep learning for brain age prediction and longitudinal tracking in children through adulthood. Imaging Neuroscience 2, 1–14. URL: https://doi.org/10.1162/imag_a_00114, doi:10.1162/imag_a_00114.

Zhao, Y., Klein, A., Castellanos, F.X., Milham, M.P., 2019. Brain age prediction: Cortical and subcortical shape covariation in the developing human brain. NeuroImage 202. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85071939061&doi=10.1016%2fj.neuroimage.2019.116149&partnerID=40&md5=b9d7beb8193d1d135757c917ea2d2423, doi:10.1016/j.neuroimage.2019.116149.

Zhou, Z., Li, Y., Zhang, Y., Liu, J., Ai, H., Liu, M., Qiu, J., Luo, Y.J., Xu, P., 2023. Differential effects of generalized anxiety and separation anxiety on brain structural development during adolescence. Journal of Affective Disorders 339, 478–485. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85165102155&doi=10.1016%2fj.jad.2023.07.056&partnerID=40&md5=c2eb74305be3269b1ca61aab1e6980f3https://www.sciencedirect.com/science/article/abs/pii/S0165032723008753?via%3Dihub, doi:10.1016/j.jad.2023.07.056.

Appendix A. Glossary

Acronyms

BabyPy A developmental brain-age model for infancy, childhood, and adolescence.

BrainPAD Brain Predicted Age Difference.

MRI Magnetic Resonance Imaging.

GMV Grey Matter Volume.

WMV White Matter Volume.

sGMV subcortical Grey Matter Volume.

MAE Mean Absolute Error.

MSE Mean Squared Error.

ExT Extra Trees Regression.

SVM Support Vector Machine.

MLP Multilayer Perceptron.

LnR Linear Regression.

IQR Interquartile Range.

SD Standard Deviation.

GPR Gaussian Process Regression.

ANTs Advanced Normalisation Tools.

SyN Symmetric Diffeomorphic Normalisation.

TBTV Total Brain Tissue Volume.

VBM Voxel-Based Morphometry.

EEG Electroencephalography.